Nonlinear Dyn (2025) 113:4077–4096
https://doi.org/10.1007/s11071-024-10477-2

RESEARCH

Model-free event-triggered resilient control for discrete-time nonlinear systems under sparse actuator attacks via GrHDP

Pengtao Song · Qingyu Yang · Zhifen Zhang · Donghe Li

Received: 24 July 2024 / Accepted: 8 October 2024 / Published online: 4 November 2024

Abstract Nonlinear modeling, multi-uncertainty analysis and robust control are the main challenges for networked control applications in industrial scenarios. To address these difficulties, this paper proposes a resilient control scheme for nonlinear cyber-physical systems with unknown disturbances and actuator attacks. First, a novel composite control framework under the strategy of goal representation heuristic dynamic programming (GrHDP) is constructed to mitigate adverse attacks, where a three-player zero-sum game (ZSG) is formulated to solve the optimal policy pair. Then, an adaptive event-triggered mechanism with performance guarantee is designed to save communication resources and improve system resilience, which does not rely on explicit dynamics and achieves both short- and long-term response performance. Under the ZSG issue, a Nash equilibrium-based GrHDP algorithm is developed to implement such a scheme, and the weight updating rules of each network are derived respectively. Furthermore, both the convergence of the performance index and the uniform ultimate boundedness of the closed-loop system are analyzed rigorously. Finally, the effectiveness of the proposed method is verified by an aero-engine networked system.

Keywords Cyber-physical systems · Actuator attacks · Event-triggered control · Goal representation heuristic dynamic programming (GrHDP) · Supplementary control

P. Song · Q. Yang · D. Li
School of Automation Science and Engineering, Xi'an Jiaotong University, Xi'an 710049, China

Q. Yang (✉)
State Key Laboratory for Manufacturing System Engineering, Xi'an Jiaotong University, Xi'an 710049, China
e-mail: yangqingyu@mail.xjtu.edu.cn

Z. Zhang
School of Mechanical Engineering, Xi'an Jiaotong University, Xi'an 710049, China

1 Introduction

Cyber-physical systems (CPSs), integrating the merits of computation, communication and control, have become the key to intelligent evolution in industrial scenarios, including multi-agent system [1], power system [2,3], electromechanical system [4], etc. With the increasing scale and complexity of CPSs, higher requirements are placed on system stability and robustness. Specifically, the robustness issue is manifested in two aspects. On the one hand, multiple uncertainties are ubiquitous in physical plant, such as system nonlinearity and external noise, making traditional control methods unable to meet the ever-increasing requirements of practical engineering [5]. On the other hand, with the involvement of communication networks, the system is exposed to open and shared environments and more vulnerable to malicious attacks that can degrade system performance and, in severe cases, cause catastrophic consequences [6]. Over the past decade, scholars in both academia and industry have devoted exten-

sive attention to this issue and conducted numerous meaningful works [7–9].

Significantly, plenty of research has focused on uncertainty modeling and robust control of complex CPSs, especially for network-induced factors, such as delay [10], scheduling [11,12], attack [6,13], etc. Since most uncertainties can be characterized as bounded disturbances, feedback control with H_∞ technology has been widely used for anti-disturbance control [2,14]. They are mainly designed by Lyapunov functions and linear matrix inequalities (LMIs), so the closed-loop system may suffer from high conservatism and low robustness. Although several advanced schemes have been implemented to improve system resilience, such as sliding mode control (SMC) [4,15], model predictive control (MPC) [16], neural network control [17], etc., they extremely rely on explicit system models and some strict assumptions, e.g., matched disturbance, Gaussian noise, or polyhedral uncertainty, greatly restricting their engineering applications. Moreover, most of these results fail to tackle optimality, which is critical for some high-performance scenarios, such as steady-state improvement for DC microgrids [18], power optimization for energy systems [19], and cost optimization for multi-agent systems [20].

As an iterative optimization algorithm, adaptive dynamic programming (ADP) provides a pathway for optimal control of complex physical systems by solving the Hamilton-Jacobi-Bellman (HJB) equation [21]. Significantly, there have been great advances in ADP algorithms, e.g., heuristic dynamic programming (HDP) [22], dual heuristic programming (DHP) [23], and globalized DHP [24]. Under these frameworks, policy iteration (PI) or value iteration (VI) is performed to solve the HJB equation, and neural networks (NNs) with actor-critic (A-C) structures are used to approach the optimal policy. In [24,25], the PI and VI have been developed respectively for linear or nonlinear continuous-time systems, and the convergence of algorithms is analyzed rigorously. Considering unknown and constrained dynamics, a model-free ADP algorithm was proposed and both optimal state and output feedback controllers were designed in [26]. Nonetheless, in industrial scenarios, since more digital processors are used for signal interaction, the discrete dynamics play a key role in control synthesis, and related works can be found in [24,27,28]. However, these schemes mainly rely on a discrete or fixed reward (e.g., $x_k^T Q x_k + u_k^T R u_k$) and fail to balance the short- and long-term performance. To alleviate the constraint of learning costate, the goal representation HDP (GrHDP) technique has been developed in [29], and such a scheme has been successfully applied to deal with the complex control of industrial devices [30,31].

Note that the above-mentioned methods rely on the assumption that the system information is complete and uncorrupted, as CPSs work in an open environment, malicious attacks are unavoidable. In practice, network attacks can be divided into *denial-of-service* (DoS) [2] and *deception attack* [13]. Among them, deception attacks are more sophisticated as they are carefully designed with system state and difficult to identify by measuring instruments. Recently, extensive studies have been made for the secure control of CPSs under deception attacks. For example, the optimal attack for linear quadratic Gaussian systems was studied in [32]. The authors in [33] designed an adaptive fuzzy scheme to mitigate the effects of cyberattacks. In [34], the secure estimation problem under linear deception attack was investigated and the performance upper bound was further analyzed. Unlike previous model-dependent frameworks, data-driven attack detection and control have been developed in [35,36], where diverse reinforcement learning algorithms are designed to balance the cost of attacker and defender. However, these schemes are mainly based on machine learning to implement detection and compensation, which cannot react quickly to the sudden changes of adverse attacks.

As an effective way to increase the resilience of CPSs, event-triggered (ET) schemes have gained much attention in reducing network traffic [2,37–40]. Specifically, in [37], the decentralized periodic ET scheme was designed for wireless CPSs with distributed sensors. In [38], the ET control of the crane bridge system modeled by partial differential equations (PDEs) was investigated, where an adaptive ET condition was proposed to reduce resource consumption. Considering the effect of malicious attacks, the problem of secure ET control for industrial CPSs under limited resource budget and deception attacks was addressed in [39], where a neural network-based controller was designed to ensure robustness. However, these pre-designed triggering conditions rely on explicit system dynamics and are difficult to migrate directly to the model-unknown complex system. To avoid the dependence on system dynamics, model-free ET mechanisms have been successively proposed in [41–43]. Note that among these

results, the controller is usually solved by NNs, so the reliability of the closed-loop system is hard to guarantee. In addition, the traditional proportional integral differential (PID) controller has always been in a commanding position in industrial control benefiting from its superiority in parameter tuning and strong robustness. As such, naturally emerged challenge is how to improve the resilience of CPSs while utilizing the merits of PID controller, which motivates this work.

To summarize the discussions made so far, we find that the resilient control of nonlinear CPSs is still a fairly open topic and some issues still exist to be addressed: (1) most advanced schemes for nonlinear CPSs rely heavily on explicit system models and strict assumptions to deal with multiple uncertainties [15], or train virtual controllers purely based on NNs, making them difficult to ensure diverse control requirements; (2) as deceptive attacks are usually designed arbitrarily and change suddenly [36], existing detection and compensation-based schemes struggle to migrate the impact of adverse attacks and fail to balance system cost; (3) most existing ET schemes rely on system dynamics or predefined structures, and ignore the exploration of the effect of ET conditions on system performance, and thus fail to apply to model-free scenarios.

Against these drawbacks, this paper proposes a resilient controller with adaptive ET mechanism for nonlinear CPSs under sparse actuator attack and external disturbance. The main contributions are as follows.

- A composite control framework is proposed for CPSs with actuator attacks, control saturation and external disturbances via zero-sum game (ZSG) and supplementary control. Compared with [32,33,35], this framework is model-free and combines the merits of PID controller and virtual controller trained by NNs to achieve robustness and reliability.
- An adaptive ET mechanism is developed to mitigate the adverse effect of attacks and save network bandwidth. The proposed triggering condition overcomes the dependence on explicit dynamics [2,37,40,41], and obtains considerable performance in the whole phase. Moreover, the stability and performance upper bound of the closed-loop system are analyzed rigorously.
- In the ZSG framework, a VI algorithm based on Nash equilibrium and GrHDP is constructed to implement such a scheme, where the weight updating rules of each network are derived respectively. Then, both theoretical and experimental simulations are performed to demonstrate the designed algorithm is convergent and the system state is ultimately uniformly bounded (UUB).

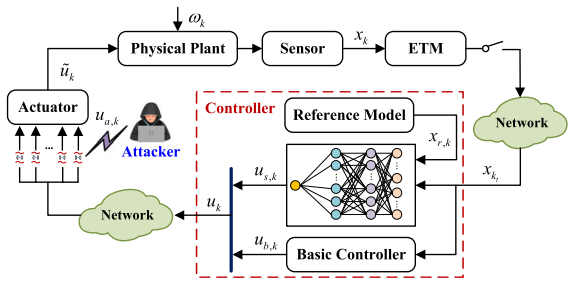

Fig. 1 Control framework of CPSs under actuator attacks

The remainder of this article is organized as follows. Section II presents the problem formulation and some preliminary results. The design of mode-free ET mechanism is developed in Section III. The ZSG problem under GrHDP framework is addressed in Sections IV. The experiment results and analysis are given in Section V. Section VI concludes this article.

Notations: Throughout this paper, $\mathbb{R}^n$ is the n dimensional Euclidean space. $\mathbb{E}\{x\}$ denotes the expectation of the stochastic variable x. The superscripts 'T' and '-1' are the matrix transpose and inverse, respectively. $\inf\{x\}$ denotes the infimum of the variable x. diag$\{...\}$ indicates the diagonal matrix. $\delta(k-i)$ is the Kronecke function that takes 1 when $k=i$ and 0 otherwise.

2 Problem formulation and preliminaries

The system architecture considered in this paper is shown in Fig. 1, in which both sensor-to-controller and controller-to-actuator channels are connected by open and shared networks. The adversary can monitor the system behavior and launch sparse actuator attacks. In particular, the system dynamics is unknown and subject to external disturbance and control saturation.

2.1 Physical system description

Consider the following discrete-time nonlinear dynamics without any prior:

$$x_{k+1} = \mathcal{F}(x_k, u_k, \omega_k) \tag{1}$$

where $x_k \in \mathbb{R}^{n_x}$ is the state vector, $u_k \in \mathbb{R}^{n_u}$ is the control input to be designed, $\omega_k \in \mathbb{R}^{n_\omega}$ is the disturbance, and k denotes the discrete-time index. $\mathcal{F}(\cdot)$ is a Lipschitz continuous function with $\mathcal{F}(0, 0, 0) = 0$. The system (1) is controllable without any constraints beyond that.

To facilitate the control synthesis, Takagi-Sugeno (T-S) fuzzy technique [2] is introduced to describe the nominal characteristics of the system (1), which is given by

Fuzzy rule i: IF $g_1(k)$ is F_1^i, $g_2(k)$ is F_2^i,..., and $g_\ell(k)$ is F_ℓ^i, THEN

$$x_{r,k+1} = A_i x_{r,k} + B_i u_{r,k} \tag{2}$$

where $i \in \{1, 2, ..., r\}$ is the number of rules, $\left\{F_j^i\right\}_{j=1}^{\ell}$ is the fuzzy set, and $g(k) = [g_1(k), g_2(k), ..., g_\ell(k)]^T$ denotes the corresponding premise variables. $x_{r,k}$ and $u_{r,k}$ are the state and control signals of the reference model, respectively. A_i and B_i are constant matrices with appropriate dimensions.

Note that the reference model does not contain multiple uncertainties and there are no strict limits on its accuracy. Then, the global T-S fuzzy model can be expressed by the following compact form

$$x_{r,k+1} = \sum_{i=1}^{r} h_i(g(k))(A_i x_{r,k} + B_i u_{r,k}) \tag{3}$$

with

$$h_i(g(k)) = \frac{\prod_{j=1}^{\ell} \mu_{ij}(g_j(k))}{\sum_{i=1}^{r} \prod_{j=1}^{\ell} \mu_{ij}(g_j(k))}, \quad \sum_{i=1}^{r} h_i(g(k)) = 1$$

where $\mu_{ij}(g_j(k))$ refers to the grade of membership of $g_j(x_k)$ in μ_{ij}, and $h_i(g(k))$ denotes the normalized membership function. Based on the parallel distributed compensation (PDC) technique [44], it is easy to design an effective controller with guaranteed performance for such a reference system.

Remark 1 The reference model can be seen as the non-necessary prior knowledge of system dynamics. It can roughly describe the nonlinear characteristics of the system (1) and provide response information to guide the control synthesis of the original system. When it's unknown, a single virtual controller can be designed with the measurement data as in [42]. Here, we assume the reference model (2) is known since it is easy to construct with the T-S fuzzy technique.

2.2 Sparse actuator attack model

In CPSs, since most actuators and communication channels are protected and the attack power is limited, deceptive attacks are usually assumed to be sparse, which is reflected in the number and frequency of attacked nodes [35]. Thus, the control signal on the actuator side is given by

$$\tilde{u}_k = u_k + \alpha_k \Gamma_k u_{a,k} \tag{4}$$

where u_k, $u_{a,k}$ and $\tilde{u}_k$ denote the well-designed control signal, false injection attack, and compromised control input, respectively. Γ_k is the attack distribution matrix to determine the attacked nodes. This model is built on the basis of [36], and we further extend it to account for the effect of attack frequency.

Define $S_{a,k} \subseteq N_a = \{1, 2, ..., n_u\}$ as the set of actuators selected by the attacker at step k. Then, the attack distribution matrix Γ_k can be indicated as $\Gamma_k = \text{diag}\{\bar{\delta}_{\sigma_k}^1, \bar{\delta}_{\sigma_k}^2, ..., \bar{\delta}_{\sigma_k}^{n_u}\}$ with $\sigma_k \in S_{a,k}$, where $\bar{\delta}_{\sigma_k}^i = \delta(\sigma_k - i)$ is the Kronecke function. $\bar{\delta}_{\sigma_k}^i = 1$ means that the i-th actuator is attacked, otherwise 0. The random variable α_k inscribes whether or not an injection attack occurs at time k and satisfies

$$\text{Prob}(\alpha_k = 1) = \bar{\alpha} \tag{5}$$

where $\alpha_k = 1$ denotes a successful attack.

For a smart adversary, the number of attacks is limited and the attack matrix Γ_k is carefully designed [6]. Furthermore, the energy of the attack signal $u_{a,k}$ is unknown but bounded to avoid being easily detected by the defender. Based on this prior, the following two types of attack patterns are considered [36]

(1) *State-independent attacks*

If $u_{a,k}$ is time-invariant, let $u_{a,k} = d$ with d being a constant. If $u_{a,k}$ is time-varying, $u_{a,k} = d_k \leq \bar{d}$ with $\bar{d}$ being the upper bounded.

(2) *State-dependent attacks*

If $u_{a,k}$ is time-invariant, let $u_{a,k} = dx_k$ with d being a constant, otherwise $u_{a,k} = d_k x_k$ with d_k being a bounded function.

The injection attack occurs randomly and fails to detect beforehand. It degrades the system performance manifested in two ways: tampering with control signals and aggravating control saturation. To improve the robustness and resilience of the system, the following composite controller is designed

$$u_k = u_{b,k} + u_{s,k} \tag{6}$$

where $u_{b,k}$ is a basic controller and can be built in advance with arbitrary structures. $u_{s,k}$ is the supplement controller to be designed.

Remark 2 In this paper, we do not design detection mechanisms to identify or reconstruct false signals as control compensation since the attack is mutable and tricky [36]. In particular, the ZSG problem is formulated in the subsequent section to solve the worst-case attack matrix Γ_k and provide guidance for the design of resilient controller. Considering that a single controller trained by NNs is unreliable and inefficient [31], a basic controller $u_{b,k}$ is introduced to maintain system stability, and a virtual controller $u_{s,k}$ is designed to further optimize the response performance of the system against external attacks.

2.3 Reward function design

Since the ideal strategy $(u_{s,k}^*, \Gamma_k^*, \omega_k^*)$ is unknown, a three-player ZSG is presented to solve the optimal policy pair, in which Player II and Player III maximize the degradation of system performance by imposing actuator attacks and deceptive disturbances, respectively, while Player I aims to search for optimal control actions to improve robustness.

Let $\mathcal{U}_s = \{u_{s,k}\}_{k=0}^{\infty}$, $\Gamma_a = \{\Gamma_k\}_{k=0}^{\infty}$, and $\mathcal{W} = \{\omega_k\}_{k=0}^{\infty}$ denote the policy spaces of three players, respectively. For the reference model, its performance will be guaranteed with a well-designed controller. If the state x_k can be accurately tracked to $x_{r,k}$, the stability of system (1) will be guaranteed. Under the framework of GrHDP [29], define the external reinforcement signal $G_k : \mathcal{U}_s \times \Gamma_a \times \mathcal{W} \to \mathbb{R}$ as

$$G_k = \mathbb{E}\left\{\bar{x}_k^T Q \bar{x}_k + \mathcal{S}(\tilde{u}_k) + \mathcal{M}(\omega_k)\right\} \tag{7}$$

where $\mathcal{M}(\omega_k) = -\gamma^2 \omega_k^T \omega_k$. $\bar{x}_k$ is the tracking error with $\bar{x}_k = x_k - x_{r,k}$. Q is a positive definite matrix. γ is a given constant to represent the attenuation level. Due to the saturation constraint, the non-quadratic power loss is defined as

$$\begin{aligned}\mathbb{E}\{\mathcal{S}(\tilde{u}_k)\} &= \mathbb{E}\left\{\int_0^{\tilde{u}_k} (\psi^{-1}(\tau/\bar{U}))^T \bar{U} R d\tau\right\} \\ &= \bar{\alpha} \int_0^{u_k + \Gamma_k u_{a,k}} (\psi^{-1}(\tau/\bar{U}))^T \bar{U} R d\tau \\ &\quad + (1-\bar{\alpha}) \int_0^{u_k} (\psi^{-1}(\tau/\bar{U}))^T \bar{U} R d\tau\end{aligned} \tag{8}$$

where R is the symmetric positive matrix, $\bar{U}$ is the upper bound of saturation, and $\psi(\cdot)$ is a bounded surjective function belonging to $\mathcal{C}^q$ $(q \geq 1)$ and $\mathcal{L}_2(\Omega)$. Without loss of generality, the tangent function $\psi(\cdot) = \tanh(\cdot)$ is used to ensure the monotonic increment of $\mathcal{S}(\tilde{u}_k)$ [42]. The partial derivatives of $\mathcal{S}(\tilde{u}_k)$ with respect to $u_{s,k}$ and Γ_k are denoted as

$$\begin{cases} \dfrac{\partial \mathcal{S}(\tilde{u}_k)}{\partial u_{s,k}} = \bar{\alpha} \bar{U} R \psi^{-1}\left(\dfrac{u_k + \Gamma_k u_{a,k}}{\bar{U}}\right) + (1-\bar{\alpha}) \bar{U} R \psi^{-1}\left(\dfrac{u_k}{\bar{U}}\right) \\ \dfrac{\partial \mathcal{S}(\tilde{u}_k)}{\partial \Gamma_k} = \bar{\alpha} \bar{U} R \psi^{-1}\left(\dfrac{u_k + \Gamma_k u_{a,k}}{\bar{U}}\right) u_{a,k}^T \end{cases} \tag{9}$$

Then, the internal reinforcement signal is given by

$$R_k = R(\bar{x}_k, u_k, \Gamma_k, \omega_k) = \sum_{i=k}^{\infty} G_i = G_k + R_{k+1} \tag{10}$$

Furthermore, the cost function is defined as

$$J(\bar{x}_k) = R_k + \varepsilon J(\bar{x}_{k+1}) \tag{11}$$

where $\varepsilon \in (0, 1)$ is a discount factor to balance the importance of immediate and future rewards. Based on Bellman's optimality principle [21], the optimal costate satisfies

$$J^*(\bar{x}_k) = \min_{u_{s,k}} \max_{\Gamma_k} \max_{\omega_k} \left\{R_k + \varepsilon J^*(\bar{x}_{k+1})\right\} \tag{12}$$

Then, the optimal policy pair $(u_{s,k}^*, \Gamma_k^*, \omega_k^*)$ can be derived from the Hamilton-Jacobi-Isaacs (HJI) equa-

tion to achieve the Nash equilibrium [41]. However, it is difficult to exactly solve due to its multivariate nonlinear partial derivative form [26]. In the subsequent section, we will introduce a GrHDP algorithm to obtain suboptimal solutions.

2.4 Problem formulation

Considering malicious attacks and limited bandwidth, event-triggered (ET) scheduling is adopted to coordinate network resources and improve system resilience. In this way, the control law is written as

$$u_k = \mu(\bar{x}_{k_t}),\ k \in [k_t,\ k_{t+1}) \tag{13}$$

where $\{k_t\}_{t=0}^{\infty}$ is a monotonically increasing triggering sequence with $k_0 = 0$, which is determined by

$$k_{t+1} = \inf\left\{k \,\middle|\, C_\eta(k) \geq 0,\ k > k_t\right\} \tag{14}$$

where $C_\eta(k)$ is the ET condition to be designed.

Based on the above statements, the main purposes of this article are threefold.

- Design a supplement controller to ensure the stability and robustness of the closed-loop system under the worst noise and attack strategies.
- Construct a model-free ET condition to alleviate network pressure and improve system resilience.
- Solve the ZSG problem (12) under the framework of GrHDP and analyze the performance upper bound of the closed-loop system.

Remark 3 The resilient control mainly reflects the stability, robustness and security of the system against external attacks, which is achieved in the subsequent sections by two parallel ways: optimal supplement control and adaptive ET mechanism. In particular, the designed ET scheme sends control signals only when the performance degradation condition is met, which can significantly reduce the number of data transmissions in the network and thus mitigate the risk of being attacked.

3 ET scheme design and performance analysis

In this section, a model-free ET mechanism is proposed, and the stability and performance upper bound of the closed-loop system are analyzed rigorously.

3.1 Event-triggered scheme design

Before building the ET mechanism, an optimal controller in the general time-driven case is designed and analyzed. The value function $Q(\bar{x}_k, u_k)$ is practically optimized to find a feasible policy pair $(u_{s,k}, \Gamma_k, \omega_k)$ in the training process, which is defined by

$$Q_k = Q(\bar{x}_k, u_k) = R_k + \min_{u_{k+1}} \varepsilon Q(\bar{x}_{k+1}, u_{k+1}) \tag{15}$$

The ZSG issue aims to find a saddle point solution $(u_{s,k}^*, \Gamma_k^*, \omega_k^*)$ such that the following Nash equilibrium holds

$$J(u_{s,k}^*, \Gamma_k, w_k) \leq J(u_{s,k}^*, \Gamma_k^*, w_k^*) \leq J(u_{s,k}, \Gamma_k^*, w_k^*) \tag{16}$$

For the admissible feedback control $u_{s,k}^*$, the optimal controller is given by

$$u_k^* = u_{s,k}^* + u_{b,k} \tag{17}$$

Then, the value function satisfies

$$\begin{aligned} Q^*(\bar{x}_k, u_k^*) &= R^*(\bar{x}_k, u_k^*) + \varepsilon Q^*(\bar{x}_{k+1}, u_{k+1}^*) \\ &= J(u_{s,k}^*, \Gamma_k^*, w_k^*) = J^*(\bar{x}_k) \end{aligned} \tag{18}$$

where $Q^*(\bar{x}_k, u_k^*)$, $R^*(\bar{x}_k, u_k^*)$ and $J^*(\bar{x}_k)$ are the optimal performance indices in the time-triggered case. Since (12) is a nonlinear PDE and difficult to solve directly, ADP and NNs are introduced in the subsequent section to solve the optimal policy pair.

Based on the GrHDP algorithm [30], the optimal controller u_k^* can be easily obtained. However, external attacks are often disguised in control signals to degrade system performance and fail to be passively and effectively suppressed. Event-triggered scheduling can actively reduce the number of signal transmissions and thus mitigate the risk of attacks. However, most existing ET conditions rely heavily on system dynamics or predefined structures [2,37,38,40], making them difficult to transfer directly to model-free scenarios. In addition, for the ET case, since the controller is updated only at the triggering moment, the performance index $J^*(\bar{x}_k)$ will inevitably decrease, so how to limit and ensure the system performance to be adjusted within a tolerable range is still a tricky problem.

To address these challenges, an adaptive ET mechanism is designed

$$C_\eta(k) = C_\eta(\bar{x}_k, \bar{x}_{k_t}) = \max\{C_{\eta_1}(k), C_{\eta_2}(k)\} < 0 \tag{19}$$

with $\Delta Q_k^* = Q^*(\bar{x}_k, \mu(\bar{x}_{k_t})) - Q^*(\bar{x}_k, u_k^*)$ and

$$\begin{cases} C_{\eta_1}(k) = \Delta Q_k^* - \eta_1 R^*(\bar{x}_k, \mu(\bar{x}_{k_t})) & \text{(20a)} \\ C_{\eta_2}(k) = G^*(\bar{x}_k, \mu(\bar{x}_{k_t})) - \eta_2 G^*(\bar{x}_k, u_k^*) & \text{(20b)} \end{cases}$$

where $C_{\eta_1}(k)$ reflects the extent to which the actual performance index $Q^*(\bar{x}_k, \mu(\bar{x}_{k_t}))$ at moment k deviates from the optimal performance index $Q^*(\bar{x}_k, u_k^*)$, which considers the offset of the internal reinforcement signal and reveals the error of future reward. While $C_{\eta_2}(k)$ characterizes the offset of the external reinforcement signal $G_k^*(\bar{x}_k, \mu(\bar{x}_{k_t}))$ under ET control, which reflects the bias of immediate reward. η_1 and η_2 are constants to govern the triggering threshold with $0 \le \eta_1 < 1$ and $\eta_2 > 1$. The control action is updated only when the triggering condition (19) is violated. Let $\{\mu(\bar{x}_{k_t})\}_{t=0}^{\infty}$ be the ET control sequence, then it satisfies $\{\mu(\bar{x}_{k_t})\}_{t=0}^{\infty} \subseteq \{u_k^*\}_{k=0}^{\infty}$.

Then, the variable substitution is applied to deal with the term $\max\{C_{\eta_1}(k), C_{\eta_2}(k)\}$. Define $\kappa(\iota_1, \iota_2)$ as a binary function satisfying $\kappa(\iota_1, \iota_2) = 1$ for $\iota_1 \ge \iota_2$, otherwise 0. Then, (19) can be rewritten as

$$C_\eta(k) = \kappa C_{\eta_1}(k) + (1-\kappa) C_{\eta_2}(k) < 0 \tag{21}$$

where κ is the abbreviation of $\kappa(C_{\eta_1}(k), C_{\eta_2}(k))$.

Remark 4 Note that the proposed ET mechanism is implemented after obtaining the time-driven optimal controller u_k^*, and it sacrifices some triggering flexibility for control reliability. The external reward signal comes from environmental feedback or changes in task demands, while the internal signal reflects or predicts the difference between internal and target performance. They are both directly related to the desired state of the system. The designed ET condition takes performance deviation as the threshold, which can comprehensively evaluate whether the control strategy needs to be updated and balance the control performance of the whole process.

3.2 Performance analysis

In the following, the stability criteria and the upper bound of performance degradation for closed-loop systems are analyzed and discussed, respectively.

Theorem 1 *For the nonlinear system (1) with the optimal policy pair $(u_{s,k}^*, \Gamma_k^*, \omega_k^*)$, the closed-loop system is asymptotically stable under the ET mechanism (21).*

Proof Consider the following Lyapunov function

$$V(k) = \kappa V_Q(k) + (1-\kappa) V_R(k) \tag{22}$$

with $V_Q(k) = \varepsilon^k Q^*(\bar{x}_k, u_k^*)$, $V_R(k) = R^*(\bar{x}_k, u_k^*)$. The proof is conducted by the following two cases.

Case 1: If the ET condition (21) is not violated, i.e., $C_\eta(k) < 0$ holds at $k \in (k_t, k_{t+1})$, we have

$$C_{\eta_1}(k) < 0, \ C_{\eta_2}(k) < 0 \tag{23}$$

Define $\bar{\eta}_1 = 1 - \eta_1$, according to (18) and (20a), we have

$$\begin{aligned} &\varepsilon^k C_{\eta_1}(k) = \varepsilon^k \{\Delta Q_k^* - \eta_1 R^*(\bar{x}_k, \mu(\bar{x}_{k_t}))\} \\ &= \varepsilon^k \{Q^*(\bar{x}_k, \mu(\bar{x}_{k_t})) - Q^*(\bar{x}_k, u_k^*) - \eta_1 R^*(\bar{x}_k, \mu(\bar{x}_{k_t}))\} \\ &= \varepsilon^k \{\varepsilon Q^*(\bar{x}_{k+1}, u_{k+1}^*) + \bar{\eta}_1 R^*(\bar{x}_k, \mu(\bar{x}_{k_t})) - Q^*(\bar{x}_k, u_k^*)\} \\ &= \Delta V_Q(k) + \varepsilon^k \bar{\eta}_1 R^*(\bar{x}_k, \mu(\bar{x}_{k_t})) \end{aligned} \tag{24}$$

Similarly, substituting (10) into (22) yields

$$\begin{aligned} &C_{\eta_2}(k) = G^*(\bar{x}_k, \mu(\bar{x}_{k_t})) - \eta_2 G^*(\bar{x}_k, u_k^*) \\ &= G^*(\bar{x}_k, \mu(\bar{x}_{k_t})) - \eta_2 \{R^*(\bar{x}_k, u_k^*) - R^*(\bar{x}_{k+1}, u_{k+1}^*)\} \\ &= G^*(\bar{x}_k, \mu(\bar{x}_{k_t})) + \eta_2 \Delta V_R(k) \end{aligned} \tag{25}$$

By combining (23)-(25), one has

$$\begin{cases} \Delta V_Q(k) = \varepsilon^k \left[C_{\eta_1}(k) - \bar{\eta}_1 R^*(\bar{x}_k, \mu(\bar{x}_{k_t}))\right] \le 0 \\ \Delta V_R(k) = \dfrac{1}{\eta_2} \left[C_{\eta_2}(k) - G^*(\bar{x}_k, \mu(\bar{x}_{k_t}))\right] \le 0 \end{cases} \tag{26}$$

which implies

$$\Delta V(k) = \kappa \Delta V_Q(k) + (1-\kappa) \Delta V_R(k) \le 0 \tag{27}$$

Case 2: If the triggering condition (21) is violated at $k = k_t$, the control signal will be updated immediately. Then, we have

$$\begin{aligned}\Delta V_Q(k) &= \varepsilon^{k+1} Q^*(\bar{x}_{k+1}, u^*_{k+1}) - \varepsilon^k Q^*(\bar{x}_k, u^*_k) \\ &= -\varepsilon^k R^*(\bar{x}_k, u^*_k) \le 0\end{aligned} \tag{28}$$

According to (10), one has

$$\begin{aligned}\Delta V_R(k) &= R^*(\bar{x}_{k+1}, \tilde{u}^*_{k+1}) - R^*(\bar{x}_k, u^*_k) \\ &= -G^*(\bar{x}_k, u^*_k) \le 0\end{aligned} \tag{29}$$

Based on the above discussion, the closed-loop system is asymptotically stable with the designed ET mechanism. This ends the proof. □

The following corollary reveals the performance guarantee of the controlled system.

Corollary 1 *For the nonlinear system (1) with the ET condition (21), the performance index is bounded and obeys the following properties*

(1) *The internal reward satisfies* $R(\bar{x}_k, \mu_k) \le \eta_2 R^*(\bar{x}_k, u^*_k)$.

(2) *The cost function satisfies* $J(\bar{x}_k, \mu_k) \le \frac{1}{\bar{\eta}_1} J^*(\bar{x}_k)$.

Proof Based on (26), we have

$$G^*(\bar{x}_k, \mu_k) \le -\eta_2 \Delta R^*(\bar{x}_k, u^*_k) \tag{30}$$

According to Theorem 1, the closed-loop system is asymptotically stable, so $R^*(\bar{x}_\infty, u^*_\infty) \to 0$ holds as $k \to \infty$. The internal reinforcement signal satisfies

$$\begin{aligned}R(\bar{x}_k, \mu_k) &= \sum_{m=0}^{\infty} G^*(\bar{x}_{k+m}, \mu_{k+m}) \\ &\le \sum_{m=0}^{\infty} \left\{-\eta_2 \Delta R^*(\bar{x}_{k+m}, u^*_{k+m})\right\} \\ &= \eta_2 \left(R^*(\bar{x}_k, u^*_k) - R^*(\bar{x}_\infty, u^*_\infty)\right) \\ &= \eta_2 R^*(\bar{x}_k, u^*_k)\end{aligned} \tag{31}$$

By combining (18) and (26), we have

$$\bar{\eta}_1 R^*(\bar{x}_k, \mu_k) \le Q^*(\bar{x}_k, u^*_k) - \varepsilon Q^*(\bar{x}_{k+1}, u^*_{k+1}) \tag{32}$$

For $n = 1, 2, ..., \infty$, the following inequalities hold

$$\begin{cases} \varepsilon \bar{\eta}_1 R^*(\bar{x}_{k+1}, \mu_{k+1}) \le \varepsilon \left(Q^*_{k+1} - \varepsilon Q^*_{k+2}\right) \\ \varepsilon^2 \bar{\eta}_1 R^*(\bar{x}_{k+2}, \mu_{k+2}) \le \varepsilon^2 \left(Q^*_{k+2} - \varepsilon Q^*_{k+3}\right) \\ ... \\ \varepsilon^n \bar{\eta}_1 R^*(\bar{x}_{k+n}, \mu_{k+n}) \le \varepsilon^n \left(Q^*_{k+n} - \varepsilon Q^*_{k+n+1}\right) \end{cases} \tag{33}$$

where $Q^*_{k+n} = Q^*(\bar{x}_{k+n}, u^*_{k+n})$. Summing over (33) yields

$$\bar{\eta}_1 \sum_{i=1}^{n} \varepsilon^i R^*(\bar{x}_{k+i}, \mu_{k+i}) \le Q^*_k - \varepsilon^{n+1} Q^*_{k+n+1} \tag{34}$$

Let $n \to \infty$, then the condition $Q^*(\bar{x}_\infty, u^*_\infty) \to 0$ holds. Successively, we have

$$J(\bar{x}_k, \mu_k) \le \frac{1}{\bar{\eta}_1} Q^*(\bar{x}_k, u^*_k) = \frac{1}{\bar{\eta}_1} J^*(\bar{x}_k) \tag{35}$$

This ends the proof. □

Remark 5 The above theorem reveals the stability of the closed-loop system. Compared with previous studies [2,38–40], this paper derives the actual performance index of the system under ET control and clearly exposes the relationship between trigger parameters and performance degradation, which can theoretically guide users to set an appropriate trigger threshold.

4 Algorithm design and NNs implementation

In this section, the formulated ZSG issue (12) is addressed under the GrHDP framework. The implementation architecture is shown in Fig. 2, where a goal network and a critic network are constructed to approach the performance indices R_k and Q_k, and two actor networks are designed for Player I and Player III to approximate the optimal policy pair $(u^*_{s,k}, \omega^*_k)$. Based on these estimates, linear programming technique is used to solve the optimal attack matrix Γ^*_k of Player II. Finally, the boundedness of the closed-loop system is analyzed rigorously.

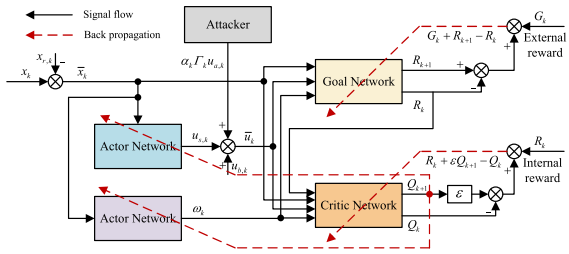

Fig. 2 Architecture of the proposed GrHDP algorithm

4.1 GrHDP network design

Note that GrHDP is an iterative optimization algorithm [29]. According to Bellman's principle, the update rule of the performance index satisfies

$$\begin{cases} R_k^{t+1} = G_k^t + R_{k+1}^t \\ Q_{k+1}^{t+1} = G_k^t + R_{k+1}^t + \varepsilon Q_{k+1}^t \end{cases} \tag{36}$$

where $G_k^t = G(\bar{x}_k, u_{s,k}^t, \Gamma_k^t, \omega_k^t)$, and t is the iteration index.

In what follows, we will give four networks to implement such an algorithm.

(1) *Goal network:* Compared with the traditional HDP algorithm [23], the internal reinforcement signal rather than external reward is designed to capture the future operation information of the system and provide guidance for the learning of costate, which is approximated by the goal network

$$\hat{R}_k^t = (w_{g2}^t)^T \phi_g(w_{g1}^T z_k) \tag{37}$$

where $(u_{s,k}, \Gamma_k)$ acts on u_k via (4), and $z_k = [\bar{x}_k^T, \tilde{u}_k^T, \omega_k^T]^T$ is the network input. w_{g1} and w_{g2} are weight matrices from the input-to-hidden (I2H) layer and hidden-to-output (H2O) layer, respectively. $\phi_g(\cdot)$ is an activation function and can be selected as the sigmoid form [42]. According to the Bellman iterative process, the error of goal network is defined as

$$e_{g,k}^t = G_k^t + \hat{R}_{k+1}^t - \hat{R}_k^t \tag{38}$$

In the learning process, to avoid excessive training parameters and improve the convergence speed, as in [31], the weight matrix w_{g1} is randomly initialized and remains unchanged. The gradient descent method [22] is used to update w_{g2} by minimizing the quadratic error $E_{g,k} = 1/2 e_{g,k}^2$, given by the following iterative law

$$w_{g2}^{t+1} = w_{g2}^t - \lambda_g \frac{\partial E_{g,k}^t}{\partial \hat{R}_k^t} \frac{\partial \hat{R}_k^t}{\partial w_{g2}^t} \tag{39}$$

where $\lambda_g > 0$ is the learning rate of the goal network.

(2) *Critic network:* The critic network is employed to approach the value function Q_k^t, which is described by

$$\hat{Q}_k^t = (w_{c2}^t)^T \phi_c(w_{c1}^T s_k) \tag{40}$$

where $s_k = [\bar{x}_k^T, \tilde{u}_k^T, \omega_k^T, R_k^T]^T$ is the input of critic network. w_{c1} and w_{c2} are bounded weight matrices from I2H layer and H2O layer, respectively. $\phi_c(\cdot)$ is the activation function. Based on (15), the error of value function is defined as

$$e_{c,k}^t = \hat{R}_k^{t+1} + \varepsilon \hat{Q}_{k+1}^t - \hat{Q}_k^{t+1} \tag{41}$$

Then, the stability-assured updating law of the weight matrix w_{c2} is obtained by minimizing the error $E_{c,k} = 1/2 e_{c,k}^2$, which is given by

$$w_{c2}^{t+1} = w_{c2}^t - \lambda_c \frac{\partial E_{c,k}^t}{\partial \hat{Q}_k^t} \frac{\partial \hat{Q}_k^t}{\partial w_{c2}^t} \tag{42}$$

where $\lambda_c > 0$ is the learning rate of the critic network.

(3) *Nash equilibrium for Player II:* The core task of Player II is to construct an optimal attack matrix Γ_k^* to launch actuator attacks. Note that the attack matrix Γ_k solved by (12) takes arbitrary distribution within the enclosed space $\Omega_\Gamma \in \mathbb{R}^{n_u \times n_u}$, which hardly meets the properties of sparse attacks [35]. Meanwhile, although both Player II and Player III optimize the cost function towards the same direction, i.e., maximally degrading the system performance, they are independent of each other and act on the system via different channels.

At each sampling moment k, the attacker randomly decides whether to launch an attack or not due to limited power. When $\alpha_k = 1$, the malicious attacker will select m nodes from all actuators to inject false data. Let $\{k_\alpha^t\}_{t=0}^\infty$ as the time sequence of successfully launched attacks. Then, the malicious attack at moment k_α^t can be expressed as $\Gamma(k_\alpha^t) = \text{diag}\{\bar{\delta}_{k_\alpha^t}^1, \bar{\delta}_{k_\alpha^t}^2, ..., \bar{\delta}_{k_\alpha^t}^{n_u}\}$, and

Algorithm 1: GrHDP for multi-player ZSG

Input: Attack fators $\bar{\alpha}$ and m, system factors x_0, Q, R, $\bar{U}$, ε and γ, iteration threshold σ and step $t_{\max}$.
Output: Optimal policy in the time-triggered case.

```
Set t = 0, R_k^0 = 0 and Q_k^0 = 0.
Construct GrHDP networks based on (37)-(50).
Initialize the weights and learning rates of networks.
while t ≤ t_max do
    Update the weight matrix w_g2 based on (39);
    Update the weight matrix w_c2 based on (42);
    Solve the optimization problem (46) with linear
    programming to obtain the attack strategy Γ_k^t;
    Update the weight matrices w_u2 and w_ω2 based on
    (49) and (52);
    Calculate the value function Q̂_k^{t+1} based on (36);
    if |Q̂_k^{t+1} − Q̂_k^t| ≤ σ then
        Solve the control law using the latest weights:
        u*_{s,k} = (w_{u2}^t)^T φ_u(w_{u1}^T x̄_k).
    else
        t = t + 1.
    end
end
return The optimal policy pair (u*_{s,k}, Γ_k^*, ω_k^*).
```

the whole ZSG problem is formulated as

$$\begin{aligned} &\min_{u_{s,k}} \max_{\Gamma_k} \max_{\omega_k} \; Q(\bar{x}_k, u_k) \\ &\text{s.t.} \; \textstyle\sum_{i=1}^{n_u} \bar{\delta}^i_{k^t_\alpha} = m \end{aligned} \tag{43}$$

According to stochastic game theory [36], the Nash equilibrium (NE) is a joint optimal policy, in which each player's strategy is optimal of the rest. Thus, the optimal attack problem is equivalent to finding a NE solution to the ZSG issue (43). Based on Bellman's principle, the optimal attack matrix satisfies

$$\begin{aligned} &\Gamma_k^* = \arg\max_{\Gamma_k} \{R(\bar{x}_k, u_k) + \varepsilon Q^*(\bar{x}_{k+1}, u_{k+1})\} \\ &\text{s.t.} \; \textstyle\sum_{i=1}^{n_u} \bar{\delta}^i_{k^t_\alpha} = m \end{aligned} \tag{44}$$

Sequentially, the optimal policy $(u^*_{s,k}, \omega^*_k)$ is computed by

$$\begin{cases} u^*_{s,k} = \bar{U}\psi(-\dfrac{1}{2}(\bar{U}R)^{-1}(\dfrac{\partial \bar{x}_{k+1}}{\partial u_{s,k}})^T \nabla Q^*_{k+1}) & (45a) \\ \omega^*_k = \dfrac{1}{2\gamma^2}\left(\dfrac{\partial \bar{x}_{k+1}}{\partial \omega_k}\right)^T \nabla Q^*_{k+1} & (45b) \end{cases}$$

Since the internal reinforcement signal R_k and optimal Q-function Q^*_k are unknown, an alternative is to approximate the problem (44) with the output of goal network and critic network. By doing so, (44) is transformed into

$$\begin{aligned} &\Gamma_k^t = \arg\max_{\Gamma_k} \{\hat{R}^t_k + \varepsilon \hat{Q}^t_{k+1}\} \\ &\text{s.t.} \; \textstyle\sum_{i=1}^{n_u} \bar{\delta}^i_{k^t_\alpha} = m \end{aligned} \tag{46}$$

The problem (46) is convex and can be solved by linear programming. After that, two actor networks are designed for Player I and Player III to get the optimal policy $(u^*_{s,k}, \omega^*_k)$.

(4) *Actor network for Player I:* The task of Player I is to obtain an optimal supplementary control law, which is approximated by

$$\hat{u}^t_{s,k} = (w^t_{u2})^T \phi_u(w^T_{u1}\bar{x}_k) \tag{47}$$

where w_{u1} and w_{u2} are the network weights of Player I, and $\phi_u(\cdot)$ is the corresponding activation function. Define the estimation error as $e^t_{u,k} = \hat{u}^t_{s,k} - u^t_{s,k}$. Considering the constraint of control saturation, according to (45a), we have

$$u^t_{s,k} = \bar{U}\psi(-\frac{1}{2}(\bar{U}R)^{-1}(\frac{\partial \bar{x}_{k+1}}{\partial u^t_{s,k}})^T \nabla \hat{Q}^t_{k+1}) \tag{48}$$

Then, the weight matrix w_{u2} is updated by minimizing the error $E_{u,k} = 1/2 e^2_{u,k}$

$$w^{t+1}_{u2} = w^t_{u2} - \lambda_u \frac{\partial E^t_{u,k}}{\partial \hat{u}^t_{s,k}} \frac{\partial \hat{u}^t_{s,k}}{\partial w^t_{u2}} \tag{49}$$

where $\lambda_u > 0$ is the learning rate.

(5) *Actor network for Player III:* The role of Player III is to find the worst disturbance in $\mathcal{L}_2[0, \infty)$ to degrade the system performance, which is given by

$$\hat{\omega}^t_k = (w^t_{\omega 2})^T \phi_\omega(w^T_{\omega 1}\bar{x}_k) \tag{50}$$

where $w_{\omega 1}$ and $w_{\omega 2}$ are the network weights of Player III, and $\phi_\omega(\cdot)$ is the bounded activation function. Based on (45b), we have

$$\omega^t_k = \frac{1}{2\gamma^2}\left(\frac{\partial \bar{x}_{k+1}}{\partial \omega^t_k}\right)^T \nabla \hat{Q}^t_{k+1} \tag{51}$$

Define the estimation error as $e^t_{\omega,k} = \hat{\omega}^t_k - \omega^t_k$ and update the weight matrix $w_{\omega 2}$ by minimizing $E_{\omega,k} =$

$1/2e_{\omega,k}^2$, one has

$$w_{\omega 2}^{t+1} = w_{\omega 2}^t - \lambda_\omega \frac{\partial E_{\omega,k}^t}{\partial \hat{\omega}_k^t} \frac{\hat{\omega}_k^t}{\partial w_{\omega 2}^t} \tag{52}$$

where $\lambda_\omega > 0$ is the learning rate of Player III.

4.2 Algorithm implementation

In the following, aiming to obtain the optimal solution of the above three-player ZSG problem, a NE-based GrHDP scheme is proposed in Algorithm 1.

Note that the designed ET condition requires the knowledge of system indices $Q^*(\bar{x}_k, u_k^*)$ and $G^*(\bar{x}_k, u_k^*)$ under time-triggered control, so in Algorithm 1, the training process of network weights does not involve ET patterns. After the training process is completed, the weights of GrHDP networks are maintained to output appropriate strategies according to the current operation state of the system. On this basis, the ET mechanism (21) is introduced into the controlled system to reduce the communication cost, and the implementation process is shown in Algorithm 1.

As the offline training and online deployment strategy is employed, the computational complexity of the proposed scheme is mainly reflected in the offline training phase of the supplement controller, which consumes some computational resources to train the neural network. While in the actual deployment phase, the computing power mainly depends on the forward propagation process of the neural network, which is linear and the resources required for each computation are fixed, so the real-time generation of control actions can be guaranteed.

4.3 Performance analysis

In this section, the convergence of both the algorithm implementation and the closed-loop system are analyzed.

Theorem 2 *For the given initial values R_k^0 and Q_k^0, when applying the following iterative process*

$$\begin{cases} u_k^t = \arg\min\limits_{u_k}\{G_k + R^t(\bar{x}_{k+1}, \\ \quad u_{k+1}^{t-1}) + \varepsilon Q^t(\bar{x}_{k+1}, u_{k+1}^{t-1})\} \\ R^{t+1}(\bar{x}_k, u_k) = G_k + R^t(\bar{x}_{k+1}, u_{k+1}^t) \\ Q^{t+1}(\bar{x}_k, u_k) = G_k + R^t(\bar{x}_{k+1}, \\ \quad u_{k+1}^t) + \varepsilon Q^t(\bar{x}_{k+1}, u_{k+1}^t) \end{cases}$$

Algorithm 2: Event-triggered control scheme

Input: Event-triggered parameters η_1 and η_2, optimal policy pair $(u_{s,k}^*, \Gamma_k^*, \omega_k^*)$ under time-triggered control, maximum step $K_{\max}$.
Output: Event-triggered control sequence.

```
while k ≤ K_max do
    Set k = 0, t = 0;
    Calculate the triggering function C_η(x̄_k, x̄_{k+1});
    if C_η(x̄_k, x̄_{k+1}) ≥ 0 then
        t = t + 1;
        Update the control law: μ(x̄_{k_t}) = û*(x̄_{k+1}).
    else
        Keep the input μ(x̄_{k_t}) = û*(x̄_k) unchanged.
    end
    k = k + 1.
end
return The event-triggered control law {μ(x̄_{k_t})}_{t=0}^∞.
```

Then, $Q^t(\bar{x}_k, u_k)$ is monotonically decreasing with t, and $Q^t(\bar{x}_k, u_k) \to Q^*(\bar{x}_k, u_k^*)$ holds when $t \to \infty$. Successively, it gives $(u_{s,k}^\infty, \Gamma_k^\infty, \omega_k^\infty) \to (u_{s,k}^*, \Gamma_k^*, \omega_k^*)$.

Proof Algorithm 1 is constructed based on Theorem 2 and employs NNs to solve the HJI equation. The above theorem reveals the convergence of Algorithm 1 and its proof is in line with [42]. The formulated ZSG is approximated by the GrHDP strategy, and then the optimal policy pair is obtained by executing the above iterations. With the help of Lyapunov theory, it is easy to prove that the estimation error of networks (37)-(52) is bounded [31], and hence omitted here. □

Based on Theorem 2, the following assumption is imposed.

Assumption 1 The performance indices $\mathcal{S}(\hat{\tilde{u}}_k)$ and $\mathcal{M}(\hat{\omega}_k)$, and the estimation errors ς_k and ζ_k are bounded, i.e.,

(1) There exist positive scalar pairs (β_{s1}, β_{s2}), $(\beta_{\omega 1}, \beta_{\omega 2})$, and (β_{r1}, β_{r2}) such that the following inequalities hold.

$$\begin{cases} \beta_{s1}\|\bar{x}_k\|^2 \le \mathcal{S}(\hat{\tilde{u}}_k) \le \beta_{s2}\|\bar{x}_k\|^2 \\ \beta_{\omega 1}\|\bar{x}_k\|^2 \le \mathcal{M}(\hat{\omega}_k) \le \beta_{\omega 2}\|\bar{x}_k\|^2 \\ \beta_{r1}\|\bar{x}_k\|^2 \le \hat{R}(\bar{x}_k, \hat{u}(k)) \le \beta_{r2}\|\bar{x}_k\|^2 \end{cases} \tag{53}$$

(2) Define the estimation errors as $\varsigma_k = Q^*(\bar{x}_k, u_k^*) - \hat{Q}(\bar{x}_k, \hat{u}_k)$ and $\zeta_k = R^*(\bar{x}_k, u_k^*) - \hat{R}(\bar{x}_k, \hat{u}_k)$. There exist two upper bounds ς_M and ζ_M such that $\|\varsigma_k\| \le \varsigma_M$ and $\|\zeta_k\| \le \zeta_M$ hold.

Theorem 3 *The system state $\bar{x}_k$ is uniformly ultimately bounded under the solved policy pair and ET mechanism.*

Proof Choose (22) as the Lyapunov function and consider the following two cases.

Case 1: If $C_\eta(k) < 0$ holds at $k \in (k_t, k_{t+1})$, one has

$$\begin{aligned}
&\Delta V_Q(k) = \varepsilon^{k+1} Q^*(\bar{x}_{k+1}, u^*_{k+1}) - \varepsilon^k Q^*(\bar{x}_k, u^*_k) \\
&= \varepsilon^{k+1} \hat{Q}(\bar{x}_{k+1}, \hat{u}_{k+1}) - \varepsilon^k \hat{Q}(\bar{x}_k, \hat{u}_k) + \Delta\bar{\varsigma}_k \\
&= \varepsilon^k [C_{\eta_1}(k) - \bar{\eta}_1 \hat{R}(\bar{x}_k, \mu(\bar{x}_{k_t}))] + \Delta\bar{\varsigma}_k \\
&\le -\varepsilon^k \bar{\eta}_1 \hat{R}(\bar{x}_k, \mu(\bar{x}_{k_t})) + \Delta\bar{\varsigma}_k \\
&\le -\varepsilon^k \bar{\eta}_1 \beta_{r1} \left\| \bar{x}_{k_t} \right\|^2 + \varsigma_M
\end{aligned} \tag{54}$$

where $\Delta\bar{\varsigma}_k = \varepsilon^{k+1}\varsigma_{k+1} - \varepsilon^k \varsigma_k$. Thus, when the system state $\bar{x}_{k_t}$ is out of the bounded domain, i.e., $\left\| \bar{x}_{k_t} \right\| \ge \sqrt{\frac{\varsigma_M}{\varepsilon^k \bar{\eta}_1 \beta_{r1}}}$, the condition $\Delta V_Q(k) \le 0$ holds such that it converges asymptotically to the equilibrium point.

Define $\Delta\zeta_k = \zeta_{k+1} - \zeta_k$ and $\tilde{\eta}_2 = 1/\eta_2$, based on (26), we have

$$\begin{aligned}
&\Delta V_R(k) = R^*(\bar{x}_{k+1}, u^*_{k+1}) - R^*(\bar{x}_k, u^*_k) \\
&= \hat{R}(\bar{x}_{k+1}, \hat{u}_{k+1}) - \hat{R}(\bar{x}_k, \hat{u}_k) + \Delta\zeta_k \\
&= \tilde{\eta}_2 \left[C_{\eta_2}(k) - \hat{G}(\bar{x}_k, \mu(\bar{x}_{k_t}) \right] + \Delta\zeta_k \\
&\le \tilde{\eta}_2 \left(\bar{x}_k^T Q \bar{x}_k + S(\hat{\bar{u}}_{k_t}) + M(\hat{\omega}_{k_t}) \right) + \Delta\zeta_k \\
&\le \tilde{\eta}_2 \left[\lambda_{\min}(Q) \|\bar{x}_k\|^2 + (\beta_{s1} + \beta_{\omega 1}) \left\| \bar{x}_{k_t} \right\|^2 \right] + \zeta_M
\end{aligned} \tag{55}$$

Similarly, $\Delta V_R(k) < 0$ holds when the following inequality is true.

$$\|\bar{x}_k\| \ge \sqrt{\frac{\eta_2 \zeta_M}{\lambda_{\min}(Q)}} \quad \text{or} \quad \left\| \bar{x}_{k_t} \right\| \ge \sqrt{\frac{\eta_2 \zeta_M}{\beta_{s1} + \beta_{\omega 1}}} \tag{56}$$

Thus, when the ET condition is not violated at $k \in (k_t, k_{t+1})$, the system state $\bar{x}_k$ is UUB.

Case 2: If $C_\eta(k) > 0$ holds at $k \in (k_t, k_{t+1})$. At this time, the update law of the control signal is the same as the time-triggered mechanism. Then, based on Theorem 1, the system is asymptotically stable. This ends the proof. □

Remark 6 The above theorem reveals the boundedness of the closed-loop system under the proposed scheme. Note that the policy pair solved by Algorithm 1 is suboptimal due to the presence of network estimation errors [31]. But the basic controller $u_{b,k}$ can provide an additional way to further deal with multiple uncertainties and improve system robustness.

4.4 Controller parameter analysis

There are three types of parameters need to be configured to achieve the resilient control scheme.

(1) *Performance parameters:* The performance parameters $(Q, R, \gamma, \varepsilon)$ are mainly used to establish the optimization criteria and balance the control power, steady-state error and anti-disturbance performance. They are often developed by the user based on actual control requirements.

(2) *Network training parameters:* The learning rate parameters $(\lambda_g, \lambda_c, \lambda_u, \lambda_\omega)$ of neural networks affect the training performance and convergence speed of the model, and users can choose these parameters according to the network complexity and dataset size.

(3) *ET parameters:* The ET parameters η_1 and η_2 are employed to balance the network quality of service (QoS) and quality of control (QoC). As the trigger parameters η_1 and η_2 increase, the number of triggers becomes less, the resource is saved more, the performance index of the system becomes larger, and the control performance worsens. Corollary 1 also reveals the relationship between the ET parameters and the deterioration of system performance indices, which can also provide guidance to users.

5 Simulation results and discussions

In this section, simulation validations are conducted to verify the effectiveness of the proposed scheme.

5.1 System setup

This section gives a simulation example on an aero-engine networked system operating in an open environment, and the system is subject to external disturbances and cyberattacks. Figure 3 depicts the structure of the simulation platform and it is consistent with [44], where TrueTime 2.0 toolbox is used to characterize the communication network. Four subsystems, including low pressure compressor (LPC) subsystem, high pressure

Table 1 Simulation-related parameters

Parameters	Q	R	γ	ε	η_1	η_2
Values	diag{2, 2}	0.01diag{1, 1}	1.2	0.8	0.1	1.05
Parameters	λ_g	λ_c	λ_u	λ_ω	σ	$t_{\max}$
Values	0.005	0.01	0.001	0.001	10^{-3}	100

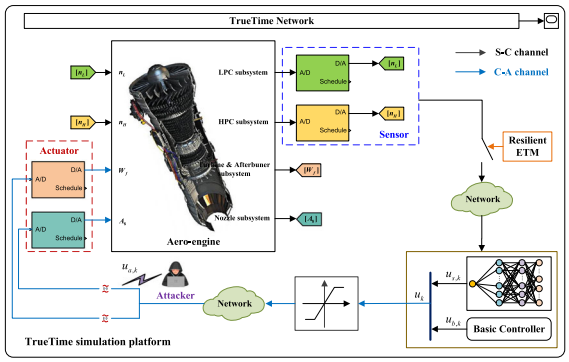

Fig. 3 TrueTime simulation framework of aero-engine networked system

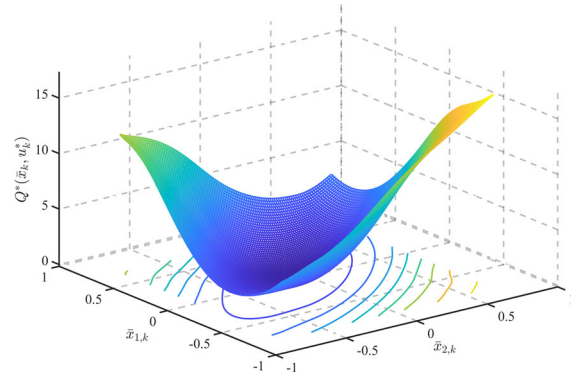

Fig. 4 Optimal value function $Q^*(\bar{x}_k, u_k^*)$ under different system states

compressor (HPC) subsystem, turbine and afterburner subsystem, and nozzle subsystem, are considered in the speed control loop. The communication nodes are implemented by TrueTime Kernel. Additionally, the composite controller and real-time ET mechanism are encoded into the intelligent sensor and controller module.

Because of the complex aerothermodynamic relationships between different components and large flight envelopes, there are no explicit nonlinear dynamics for aero-engines. Zhou et al. [45] established a T-S fuzzy model from AL-31F aero-engine, which is employed in this paper as a nominal reference system to simulate the full-envelope characteristics. The model structure is as follows:

Fuzzy rule 1: If H_f is 0km and Ma is about 0.11, then

$$A_1 = \begin{bmatrix} -1.26364 & 1.03610 \\ -0.15054 & -1.24944 \end{bmatrix},$$
$$B_1 = \begin{bmatrix} 0.35420 & 0.54739 \\ 0.35765 & 0.32159 \end{bmatrix}$$

Fuzzy rule 2: If H_f is 6.86km and Ma is about 0.75, then

$$A_2 = \begin{bmatrix} -0.68796 & 0.56623 \\ -0.10917 & -0.74529 \end{bmatrix},$$
$$B_2 = \begin{bmatrix} 0.23334 & 0.32494 \\ 0.21948 & 0.21712 \end{bmatrix}$$

Fuzzy rule 3: If H_f is 13.61km and Ma is about 1.40, then

$$A_3 = \begin{bmatrix} -0.70638 & 0.33246 \\ -0.13549 & -0.37354 \end{bmatrix},$$
$$B_3 = \begin{bmatrix} 0.10421 & 0.15536 \\ 0.10228 & 0.08665 \end{bmatrix}$$

Fuzzy rule 4: If H_f is 16.27km and Ma is about 1.44, then

$$A_4 = \begin{bmatrix} -0.89628 & 0.51077 \\ -0.18259 & -0.56626 \end{bmatrix},$$
$$B_4 = \begin{bmatrix} 0.14051 & 0.27368 \\ 0.15398 & 0.13103 \end{bmatrix}$$

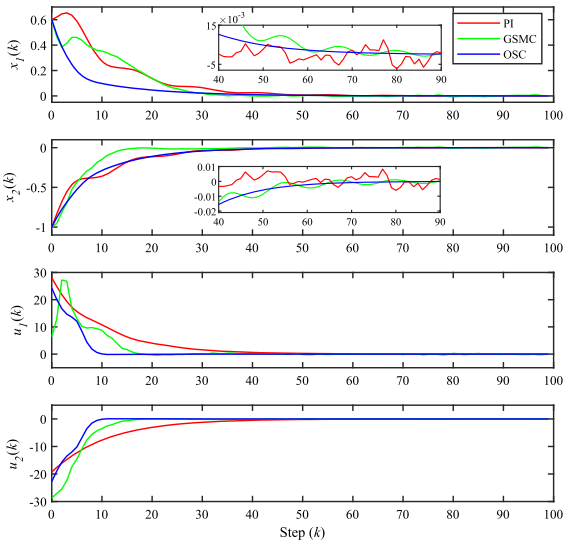

Fig. 5 Response curves of closed-loop system under different controllers

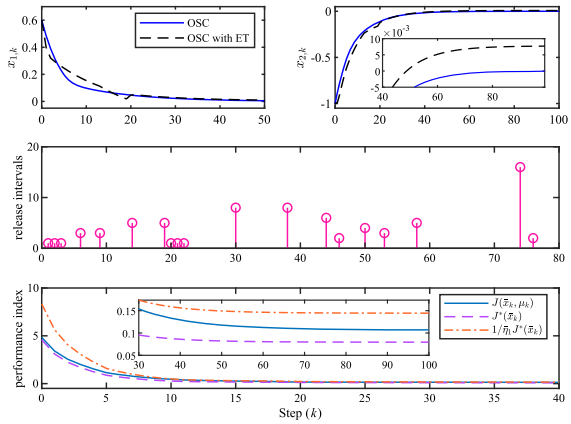

Fig. 6 System performance under different triggering conditions

In this paper, we mainly focus on speed control, so the output matrices are $C = \text{diag}\{1, 1\}$ and $D = 0$. The sampling period is set to $h = 20$ms. The saturation constraint is $|\bar{U}| = 30$. The control inputs are the fuel mass flow W_f (kg/s) and throat area in nozzle A_8 (m^2), and the outputs are the low pressure rotor speed n_L (r/min) and high-pressure rotor speed n_H (r/min). The external disturbance is random electromagnetic noise perturbed around 20% of the equilibrium point.

Note that the common scheme for aero-engine networked systems is to design feedback or PID controllers. In [44], a robust fuzzy PI scheme is designed for aero-engine systems with network-induced factors, and it is taken for comparison. Moreover, Ren et al. [46] explored a new achievement for generalized SMC (GSMC), which is more universal and can be used as a benchmark. The proposed scheme involves three types of parameters: performance parameters (Q, R, γ, ε), training parameters (λ_g, λ_c, λ_u, λ_ω, σ, $t_{\max}$) and design parameters (η_1, η_2), and their initial values are shown in Table 1. Specifically, the performance parameters and training parameters are generalized parameters set by users based on expected quadratic performance and network training loss. The design parameters are employed to balance the network quality of service (QoS) and quality of control (QoC), and users can set them according to Corollary 1 and specific control requirements. To give a clear discussion, a series of experiments are carried out to verify the control performance of the proposed scheme.

5.2 Optimal supplementary control without attack

In the first case, consider the aero-engine operating at ($H_f = 3.98, Ma = 0.38$). Set the initial state $x_0 = [0.6, -1]^T$ and $\bar{x}_0 = [0.4, -0.2]^T$. When there are no actuator attacks, i.e., $\bar{\alpha} = 0$, under the framework of GrHDP, the structures of NNs for goal network, critic network, and actor networks are chosen as $5-32-1$, $6-32-1$, $2-20-2$, and $2-20-1$, respectively. The initial weights for each network are randomly valued within $[-1, 1]$. Define the state exploration space as $\mathcal{X} = \{\bar{x}_k \,|-0.8 \le \bar{x}_{1,k} \le 0.8, -0.8 \le \bar{x}_{2,k} \le 0.8\}$. We uniformly select 40,000 groups of states to obtain the optimal value function $Q^*(\bar{x}_k, u_k^*)$. The GrHDP algorithm executes no more than $t_{\max}$ to achieve the calculation accuracy σ. Figure 4 depicts the output of the well-trained Q-function under different system states. As the system state converges to the equilibrium point, the optimal value $Q^*(\bar{x}_k, u_k^*)$ monotonically decreases towards 0, which confirms that the trained critic network is capable of evaluating the entire state space.

The response curves of the closed-loop system with different controllers are shown in Fig. 5. One can see that the PI controller is sensitive to system uncertainty and leads to a large steady-state error. Although GSMC has good anti-disturbance capability, it requires high gains for robustness and easily causes a wide quasi-sliding band. Thus, some additional measures such as disturbance compensation or adaptive switching mech-

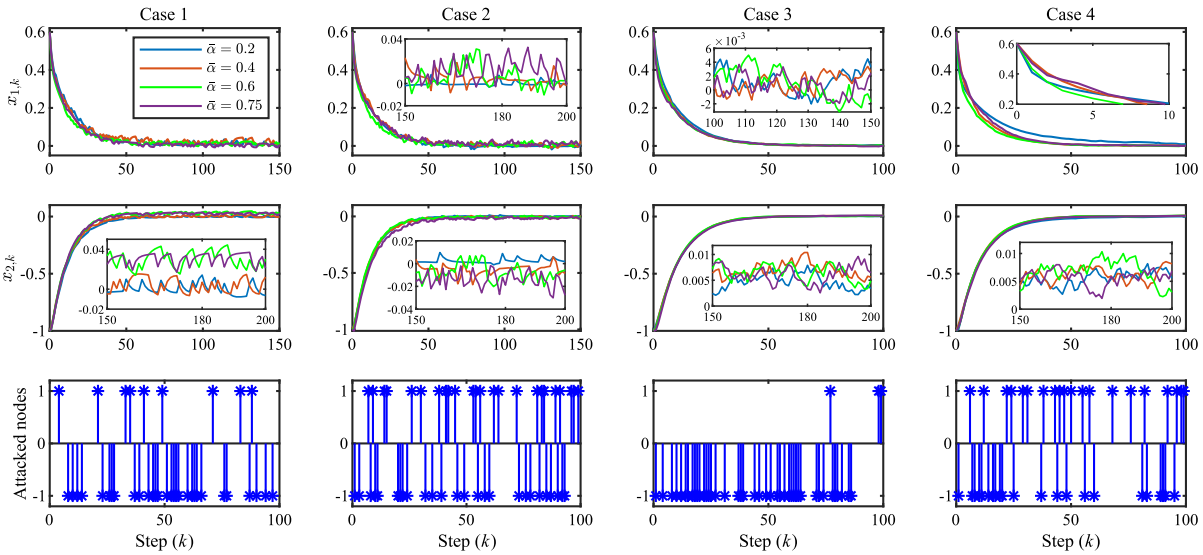

Fig. 7 Response curves of the closed-loop system under different attack frequency $\bar{\alpha}$ and mode (Case 1/Case 3: state-independent/dependent constant attack; Case 2/Case 4: state-independent/dependent time-varying attack)

anisms need to be introduced to alleviate the chattering phenomenon, which also increases the design complexity. Significantly, the proposed optimal supplementary controller (OSC) introduces an additional control action $u_{s,k}$ to fine-tune system performance and further enhances the anti-disturbance capability by imposing the disturbance generation network and ZSG problem, achieving no obvious overshoot and steady-state error of less than 1%. Figure 6 shows the response curves of the system under different triggering cases. Under the proposed ET mechanism, the communication consumption is reduced by nearly 70% with the sacrifice of tolerable control performance. Meanwhile, the actual performance index $J(\bar{x}_k, \mu_k)$ is always bounded at $[J^*(\bar{x}_k), 1/\bar{\eta}_1 J^*(\bar{x}_k)]$, which is also consistent with the conclusion of Corollary 1.

5.3 Optimal supplementary control with actuator attack

In this section, consider the aero-engine operating at the afterburner state with ($H_f = 0$, $Ma = 0$). As discussed in Section II, there are four types of false attack signals involved, including state-independent/dependent constant attacks [35], and state-independent/dependent time-varying attacks [36]. The initial parameters are set in line with Table 1, and the attack signal $u_{a,k}$ is chosen randomly or fixedly from $[-2, 2]$ at each attack moment, i.e. $d = \bar{d} = 2$. Figure 7 shows the response curves of the closed-loop system under different attack frequencies and modes, where Case 1-Case 4 indicate four attack types, respectively. As can be seen from Fig. 7, the system performance degrades with the attack frequency $\bar{\alpha}$, and state-independent attacks are more destructive to system performance than state-dependent attacks, showing whole-phase chattering and large steady-state errors. This is mainly because the proposed scheme can ensure the asymptotic stability of the closed-loop system, and the state-dependent attack intensity tends to zero with the convergence of state variable, thus achieving good steady-state performance.

The response curves of attack signal and disturbance signal for the four attack types at $\bar{\alpha} = 0.4$ are shown in Fig. 8. They are generated by the optimization problem (46) and actor network (50) respectively and act on the system via different transmission channels. Specifically, the disturbance signal oscillates between $0 \sim 10\%$ of the maximum state and gradually converges with the formulated ZSG. The state-independent attacks degrade system performance and impose the greatest damage to system stability at a fixed or random intensity over the whole phase, and thus are the most effective means for the adversary to launch such attacks. Set the attack frequency $\bar{\alpha} = 0.4$, the response curves of

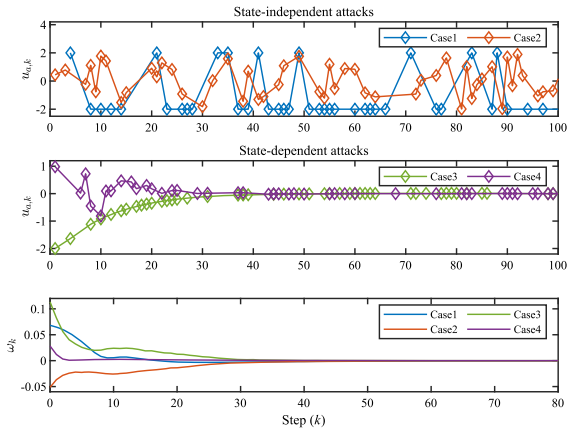

Fig. 8 Responses of attack signal and disturbance signal at $\bar{\alpha} = 0.4$

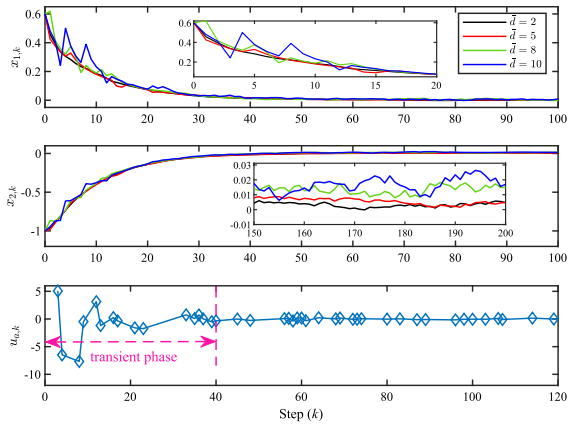

Fig. 9 Response curves of the system with different attack intensities

the system under state-dependent time-varying attacks with different attack intensities are shown in Fig. 9. Although the attack intensity has reached 30% of the saturation constraint $\bar{U}$, the proposed controller is still able to migrate the attack effect and maintain the state convergence, which also confirms the robustness of the designed composite scheme.

Figure 10 shows the system response of different controllers under the state-independent time-varying attack with $\bar{\alpha} = 0.4$ and $\bar{d} = 2$, where '1' and '-1' denote the two nodes of the actuator, respectively. At each time, the attack launched by the adversary is stochastic and near-optimal. It randomly chooses one actuator node for false data injection and acts on the system through the control channel, which can be regarded as matched sparse uncertainty but remains difficult to

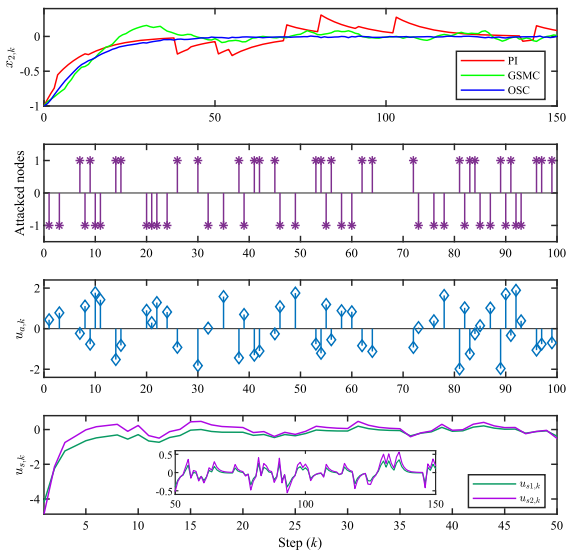

Fig. 10 Response of different controllers under state-independent attacks

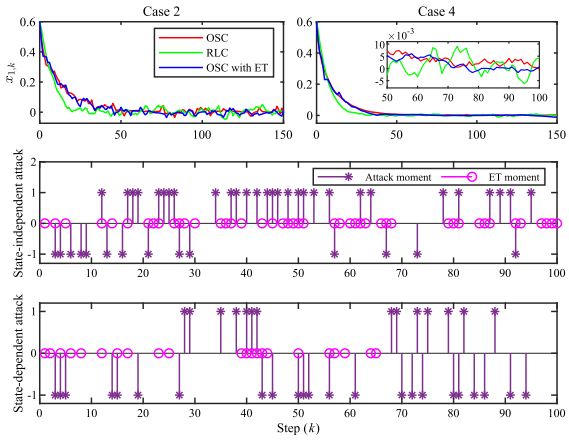

Fig. 11 System response under different triggering cases and attack modes

detect in advance. Since the construction of GSMC relies on the model prior and disturbance assumption, it is unable to effectively migrate the effects of multiple uncertainties, leading to a large overshoot and error band. Instead, the proposed scheme introduces an additional control action $u_{s,k}$ to fine-tune the system performance, which is generated by the actor network (47) and adaptively varies with the attack intensity, thus achieving strong resilience against malicious attacks.

Figure 11depicts the response curves of the system with state-independent/dependent attack modes under different triggering cases, where the mainstream rein-

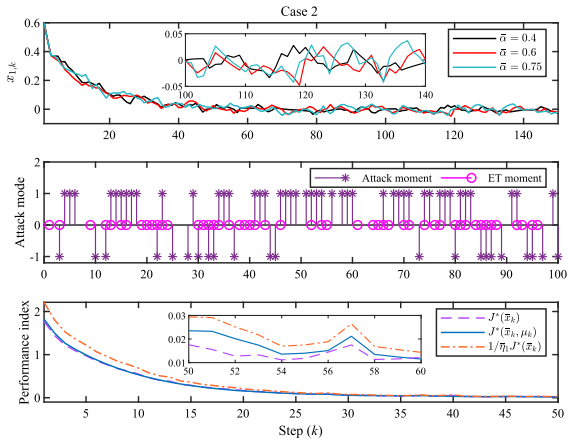

Fig. 12 Response curves of the system under ET control for different $\bar{\alpha}$

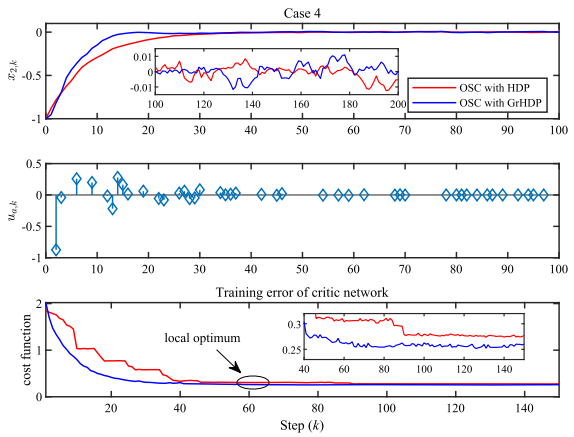

Fig. 13 Response curves of the system under different training algorithms

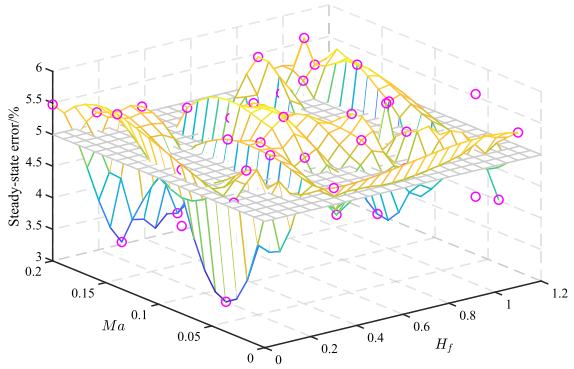

Fig. 14 Steady-state errors at different flight points with OSC with ET

forcement learning control (RLC) [36] is introduced for comparison. As shown in Fig. 11, both RLC and the proposed OSC can effectively migrate the effect of malicious attacks. However, despite its stronger nonlinear mapping ability, RLC requires more computational resources and hyperparameters to train the deep model and ignores the handling of external disturbances, which usually act directly on the system in the form of mismatched noise through other channels. With the introduction of the ET mechanism, the network sends control signals only when the ET condition (19) is satisfied, so the malicious attack only succeeds at the ET moment. In this case, the designed strategy achieves comparable performance using only about 40% of communication resources, and it can both migrate malicious actuator attacks and reduce resource consumption, which is attractive for high-frequency sampling CPSs with multiple signals.

Figure 12 displays the response curves of the system under ET control for different attack frequency $\bar{\alpha}$. One can see that the system performance deteriorates as the attack frequency $\bar{\alpha}$ increases. By introducing ET scheduling, when the attack frequency is $\bar{\alpha} = 0.75$, the adversary launches a total of 67 attacks in 100 steps, but only 36 attacks successfully act on the actuator, which confirms the resilience of the closed-loop system against malicious attacks. Meanwhile, the formulated problem (12) considers both the short- and long-term performance of the system, and the actual performance index $J(\bar{x}_k, \mu_k)$ is bounded and maintained around the near-optimal value $J^*(\bar{x}_k)$ even with frequent external attacks. Table 2 exhibits the performance index and ET number under different triggering parameters. Specifically, η_1 and η_2 focus on long-time performance and short-time performance, respectively, and the larger η_1 and η_2 are, the fewer the number of triggers, and the worse the performance of the system. Thus, users can select ET parameters according to specific control requirements to achieve the trade-off between performance and resource consumption.

Set the initial parameters consistent with Table 1 and attack frequency $\bar{\alpha} = 0.4$, the response curves of the system under different training algorithms are shown in Fig. 13, where the HDP algorithm [22] is introduced to solve the supplement controller and its value function is set to $Q_k = G_k + Q_{k+1}$, which is approximated by the critic network. As shown in Fig. 13, both HDP and GrHDP are able to solve the formulated ZSG problem. However, the reward signal of HDP usually consists of a fixed immediate reward G_k, which is difficult to adjust

Table 2 Performance index for different ET parameters

$\eta_1,\ \eta_2 = 1.05$	0	0.05	0.1	0.4
$J(\bar{x}_k, \mu_k)$	0.0098	0.0093	0.0108	0.0271
ET number	71	53	48	21
$\eta_2,\ \eta_1 = 0.1$	1	1.05	1.15	1.3
$J(\bar{x}_k, \mu_k)$	0.0089	0.0108	0.0133	0.0170
ET number	67	48	31	17

adaptively with the change of the system state [31]. Instead, GrHDP adds a goal network to HDP, which can not only receive the external reinforcement signal, but also generate the long-term reward adapted to the current state through the internal goal-oriented mechanism, so as to better regulate the mapping between input signals and output control actions, and guide the fast convergence of network training errors to avoid falling into the local optimum. Figure 14shows the response curves of the system at different flight points. One can see that the designed composite scheme is robust to flight point variations and guarantees a steady-state error of less than 6% over a wide range of flight areas. All the experimental results confirm the effectiveness of the proposed scheme.

6 Conclusion

In this article, the supplementary control of nonlinear CPSs with unknown disturbances and actuator attacks is investigated. Firstly, a novel composite control scheme under the framework of GrHDP is proposed to mitigate adverse attacks, where a three-player ZSG is formulated to solve the optimal policy pair. Then, an adaptive ET mechanism with performance guarantees is designed to save communication resources and improve system resilience. Both theoretical and experimental simulations are performed to demonstrate that the designed algorithm is convergent and the system state is UUB. The experimental results show that the constructed composite controller owns the advantages of fast convergence, small chattering, robustness to malicious, and few communication resources, which are beneficial for CPSs with high-frequency sampling and multiple uncertainties. In the future, the proposed scheme will be extended to more complex scenarios, such as dual-channel scheduling, hybrid attacks, and fault-tolerant control.

Funding This work was supported by the National Science and Technology Major Project of China under Grant J2019-V-0003-0094.

Data availability The data used to support the findings of this study are included within the article.

Declarations

Conflict of interest The authors declare that there is no Conflict of interest regarding the publication of this paper.

References

1. Wang, J., Yang, S., Wang, Q., Ji, L.: Finite-time consensus of nonlinear delayed multi-agent system via state-constraint impulsive control under switching topologies. Nonlinear Dyn. **111**(13), 12267–12281 (2023)
2. Li, X.H., Ye, D.: Security-based event-triggered fuzzy control for multiarea power systems under cross-layer DoS attacks. IEEE Trans. Circuits Syst. I, Reg. Papers **70**(7), 2995–3004 (2023)
3. Teklehaimanot, Y.K., Akingbade, F.K., Ubochi, B.C., Ale, T.O.: A review and comparative analysis of maximum power point tracking control algorithms for wind energy conversion systems. Int. J. Dyn. Control (2024). https://doi.org/10.1007/s40435-024-01434-3
4. Song, P., Yang, Q., Li, D., Zhang, Z., Peng, J.: Disturbance-compensation-based predictive sliding mode control for aero-engine networked systems with multiple uncertainties. IEEE Trans. Autom. Sci. Eng. (2024). https://doi.org/10.1109/tase.2024.3350020
5. Wu, M., Liu, L., Yu, Z.: Optimization controller synthesis using adaptive robust control Lyapunov and barrier functions for high-order nonlinear system. Nonlinear Dyn. **111**(19), 17973–17986 (2023)
6. Kim, S., Park, K.J., Lu, C.: A survey on network security for cyber-physical systems: from threats to resilient design. IEEE Commun. Surv. Tuts. **24**(3), 1534–1573 (2022)

7. Kordestani, M., Saif, M.: Observer-based attack detection and mitigation for cyber-physical systems: a review. IEEE Syst. Man, Cybern. Mag. **7**(2), 35–60 (2021)
8. Zhang, D., Li, C., Goh, H., Ahmad, T.: A comprehensive overview of modeling approaches and optimal control strategies for cyber-physical resilience in power systems. Renew. Energy **189**, 1383–1406 (2022)
9. Wan, Y., Cao, J.: A brief survey of recent advances and methodologies for the security control of complex cyber-physical networks. Sensors **23**(8), 11–34 (2023)
10. Yan, H., Ma, Q., Wang, J., Huang, H.: Fractional-order time-delay feedback control for nonlinear dynamics in giant magneto strictive actuators. Nonlinear Dyn. **112**(5), 3055–3079 (2024)
11. Yang, R., Li, G., Zhu, Y., Liu, G.P.: Event-triggered control for networked predictive control systems with time delay and external disturbance. IEEE Trans. Control Netw. Syst. **10**(4), 2120–2129 (2023)
12. Wang, Y., Wang, Z., Zou, L., Dong, H.: Observer-based fuzzy PID tracking control under try-once-discard communication protocol: an affine fuzzy model approach. IEEE Trans. Fuzzy Syst. **32**(4), 2352–2365 (2024)
13. Ji, Y., Gao, Q., Liu, J.: Adaptive resilient control for cyber-physical systems against unknown injection attacks in sensor networks. Nonlinear Dyn. **111**(12), 11105–11114 (2023)
14. Tang, Z., Xiao, W., Zhang, B.: H-infinity loop shaping control of wireless power transfer system based on generalized state space averaging model. Int. J. Circ. Theor. Appl. **52**(1), 364–379 (2023)
15. Song, T., Fang, L., Zhang, Y., Shen, H.: Recursive terminal sliding mode based control of robot manipulators with a novel sliding mode disturbance observer. Nonlinear Dyn. **112**(2), 1105–1121 (2023)
16. Shorinwa, O., Schwager, M.: Distributed model predictive control via separable optimization in multiagent networks. IEEE Trans. Autom. Control **63**(1), 230–245 (2024)
17. Su, S., Ji, Y.: Fixed-time adaptive neural network synchronization control for teleoperation system with position error constraints and time-varying delay. Nonlinear Dyn. **111**(14), 13053–13072 (2023)
18. Peng, J., Fan, B., Liu, W.: Penalty-based distributed optimal control of DC microgrids with enhanced current regulation performance. IEEE Trans. Circuits Syst. I, Reg. Papers **69**(7), 3026–3036 (2022)
19. Chen, S.Y., Chang, C.H.: Optimal power flows control for home energy management with renewable energy and energy storage systems. IEEE Trans. Energy Convers. **38**(1), 218–229 (2023)
20. Zhang, Z., Shi, Y., Zhang, Z., Li, H.: Robust optimal control of uncertain discrete-time multiagent systems with digraphs. IEEE Trans. Syst., Man, Cybern., Syst. **53**(3), 861–871 (2023)
21. Liu, D., Xue, S., Zhao, B., Luo, B., Wei, Q.: Adaptive dynamic programming for control: a survey and recent advances. IEEE Trans. Syst., Man, Cybern., Syst. **51**(1), 142–160 (2021)
22. Zhao, Q., Si, J., Sun, J.: Online reinforcement learning control by direct heuristic dynamic programming: from time-driven to event-driven. IEEE Trans. Neural Netw. Learn. Syst. **33**(8), 4139–4144 (2022)
23. Ni, Z., He, H., Zhong, X., Prokhorov, D.V.: Model-free dual heuristic dynamic programming. IEEE Trans. Neural Netw. Learn. Syst. **26**(8), 1834–1839 (2015)
24. Sun, B., Kampen, E.J.: Incremental model-based global dual heuristic programming with explicit analytical calculations applied to flight control. Eng. Appl. Artif. Intell. **89**, 103425 (2020)
25. Xiao, G., Zhang, H.: Convergence analysis of value iteration adaptive dynamic programming for continuous-time nonlinear systems. IEEE Trans. Cybern. **54**(3), 1639–1649 (2024)
26. Rizvi, S.A.A., Lin, Z.: Adaptive dynamic programming for model-free global stabilization of control constrained continuous-time systems. IEEE Trans. Cybern. **52**(2), 1048–1060 (2022)
27. Zhu, G., Li, X., Sun, R., Yang, Y., Zhang, P.: Policy iteration for optimal control of discrete-time time-varying nonlinear systems. IEEE/CAA J. Autom. Sinica **10**(3), 781–791 (2023)
28. Ye, J., Bian, Y., Luo, B.: Costate-supplement ADP for model-free optimal control of discrete-time nonlinear systems. IEEE Trans. Neural Netw. Learn. Syst. **35**(1), 45–59 (2024)
29. Zhen, N., Hai, H., Jin, W., Xin, X.: Goal representation heuristic dynamic programming on maze navigation. IEEE Trans. Neural Netw. Learn. Syst. **24**(12), 2038–2050 (2013)
30. Zhong, X., He, H.: GrHDP solution for optimal consensus control of multiagent discrete-time systems. IEEE Trans. Syst., Man, Cybern., Syst. **50**(7), 2362–2374 (2020)
31. Wang, X., Ding, D., Ge, X., Han, Q.L.: Supplementary control for quantized discrete-time nonlinear systems under goal representation heuristic dynamic programming. IEEE Trans. Neural Netw. Learn. Syst. **35**(3), 3202–3214 (2024)
32. Ren, X., Yang, G.: Kullback-Leibler divergence-based optimal stealthy sensor attack against networked linear quadratic Gaussian systems. IEEE Trans. Cybern. **52**(11), 11539–11548 (2022)
33. Bi, Y., Wang, T., Qiu, J.: Adaptive decentralized finite-time Fuzzy secure control for uncertain nonlinear CPSs under deception attacks. IEEE Trans. Fuzzy Syst. **31**(8), 2568–2580 (2023)
34. Zhang, M., Sun, Y., Yang, X., Park, J.H.: Secure state estimation of singularly perturbed switched systems with output deception attack via discretized LKF approach. IEEE Trans. Circuits Syst. II, Exp. Papers **71**(2), 792–796 (2024)
35. Zhao, Z., Xu, Y., Li, Y., Zhao, Y.: Sparse actuator attack detection and identification: a data-driven approach. IEEE Trans. Cybern. **53**(6), 4054–4064 (2023)
36. Wu, C., Pan, W., Staa, R.: Deep reinforcement learning control approach to mitigating actuator attacks. Automatica **152**, 11099 (2023)
37. Fu, A., McCann, J.A.: Dynamic decentralized periodic event-triggered control for wireless cyber-physical systems. IEEE Trans. Control Syst. Technol. **29**(4), 1783–1790 (2021)
38. Tang, L., Li, Y., Liu, Y.: Adaptive event-triggered vibration control for crane bridge system with trolley. IEEE Trans. Instrum. Meas. **72**, 1–10 (2023)
39. Ma, Y., Li, Z.: Neural network-based secure event-triggered control of uncertain industrial cyber-physical systems against deception attacks. Inf. Sci. **633**, 504–516 (2023)

40. Song, P., Yang, Q., Wen, G., Zhang, Z.: Observer-based output feedback control for networked systems with dual-channel event-triggered sampling and quantization. J. Frankl. Inst. **359**(14), 7365–7392 (2022)
41. Wang, D., Hu, L., Zhao, M., Qiao, J.: Dual event-triggered constrained control through adaptive critic for discrete-time zero-sum games. IEEE Trans. Syst., Man, Cybern., Syst. **53**(3), 1584–1595 (2023)
42. Wang, J., Wang, Y., Ji, Z.: Model-free event-triggered optimal control with performance guarantees via goal representation heuristic dynamic programming. Nonlinear Dyn. **108**(4), 3711–3726 (2022)
43. Wang, X., Hua, C., Qiu, Y.: Event-triggered model-free adaptive control for nonlinear multiagent systems under jamming attacks. IEEE Trans. Neural Netw. Learn. Syst. (2024). https://doi.org/10.1109/tnnls.2023.3279144
44. Song, P., Yang, Q., Wen, G., Zhang, Z., Peng, J.: Fuzzy H_∞ robust control for T-S aero-engine systems with network-induced factors under round-robin like protocol. Aerosp. Sci. Technol. **137**, 108258 (2023)
45. Zhou, B., Xie, S., Hui, J.: H-infinity control for T-S aero-engine wireless networked system with scheduling. IEEE Access **7**, 115662–115672 (2019)
46. Ren, Q., Kao, Y., Wang, C., Xia, H., Wang, X.: New results on the generalized discrete reaching law with positive or negative decay factors. IEEE Trans. Autom. Control **67**(2), 1046–1052 (2023)